Human-In-The-Loop Person Re-Identification

Hanxiao Wang, Shaogang Gong, Xiatian Zhu, Tao Xiang
{hanxiao.wang, s.gong, xiatian.zhu, t.xiang}@qmul.ac.uk

School of EECS, Queen Mary University of London, London E1 4NS, UK

Queen Mary University of London

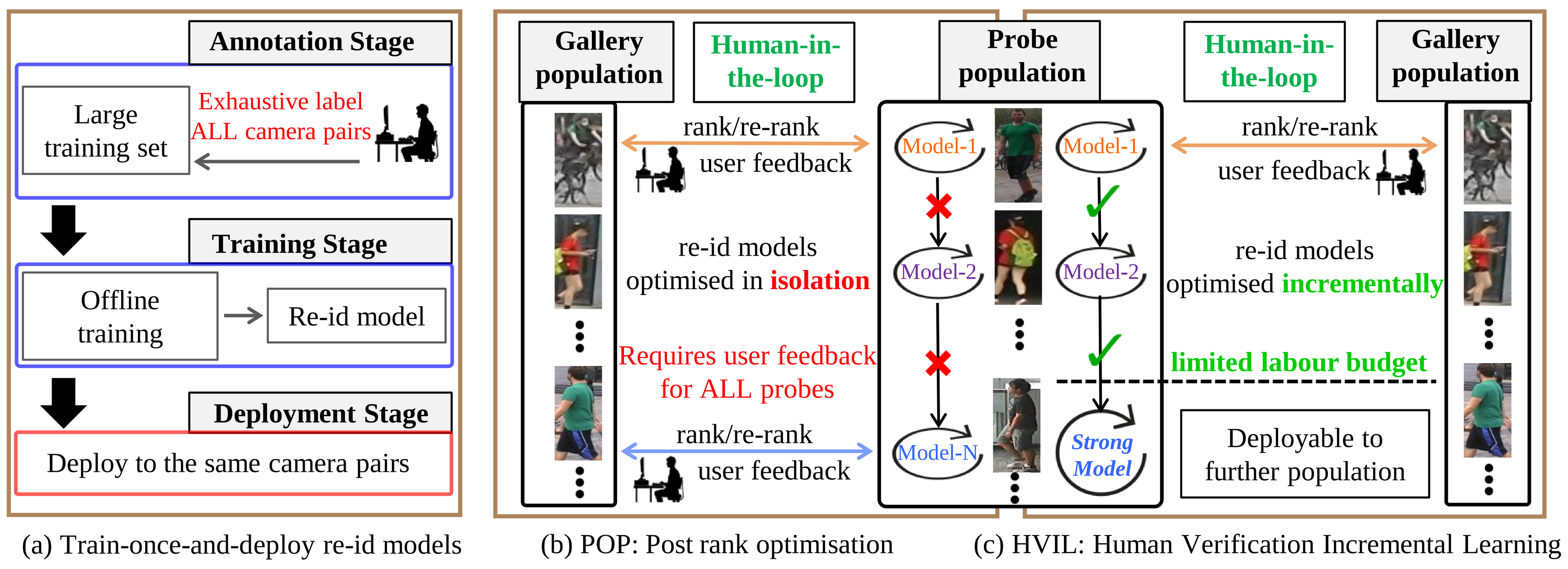
ECCV'16 EUROPEAN CONFERENCE ON COMPUTER VISION

1. Motivation

- Existing train-once-and-deploy person re-identification approaches are not scalable, and sometimes even not plausible.
  - ○ **Extensive pre-labelled training data** ○ **Small testing population**
- In a real-world, given low Rank-1 recognition rates, human operators are required to verify the unideal ranking lists.
- How to explore human-in-the-loop for person re-identification?

2. Contribution

- A Human Verification Incremental Learning (HVIL) model which enables human-in-the-loop person re-identification.
  - ○ **Less labelling effort** ○ **Flexible feedback** ○ **Immediate benefit**
- Compared to Post-rank Optimisation (POP) [1], HVIL enables incremental model improvement from cumulative human feedback.
- A Regularised Metric Ensemble Learning (RMEL) model is proposed when human feedback becomes unavailable.



(a) Train-once-and-deploy re-id models (b) POP: Post rank optimisation (c) HVIL: Human Verification Incremental Learning

Figure 1: (a) Conventional *train-once-and-deploy* re-id strategy requires pre-labelled training data collection. (b) POP [1]: A recent *human-in-the-loop re-id* approach which optimises probe-specific models in isolation. (c) HVIL: The proposed new incremental human-in-the-loop re-id model.

3. Approach

Modelling Human Feedback as a Loss Function

- An incrementally optimised ranking function, $f_{\boldsymbol{x}^p}(\boldsymbol{x}_i^g): \mathbb{R}^d \rightarrow \mathbb{R}$, $y \in L = \{m, s, w\}$ as *true-match*, *strong-negative*, and *weak-negative* respectively.

$$err(f_{\boldsymbol{x}^p}(\boldsymbol{x}^g), y) = \mathcal{L}_y(rank(f_{\boldsymbol{x}^p}(\boldsymbol{x}^g))), \quad (1)$$

- A novel re-id ranking loss is introduced:

$$\mathcal{L}_y(k) = \begin{cases} \sum_{i=1}^{k} \alpha_i, & \text{if } y \in \{m, w\} \\ \sum_{i=k+1}^{n_g} \alpha_i, & \text{if } y \in \{s\} \end{cases}, \quad \text{with} \quad \alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant 0. \quad (2)$$

- We set $\alpha_i = \frac{1}{i}$ when y indicates a *true-match*, and $\alpha_i = \frac{1}{n_g - 1}$ with n_g the gallery size when y represents a *weak-negative* or *strong-negative*.

Real-time Model Update for Instant Feedback Reward

- Consider re-id ranking model $f(\cdot)$ as a negative Mahalanobis distance:

$$f_{\boldsymbol{x}^p}(\boldsymbol{x}^g) = -\left[(\boldsymbol{x}^p - \boldsymbol{x}^g)^\top \boldsymbol{M}(\boldsymbol{x}^p - \boldsymbol{x}^g)\right], \ \boldsymbol{M} \in S_+^d. \quad (3)$$

- For real-time feedback and reward, $f(\cdot)$ is estimated on each human feedback in an online manner [2].

$$\boldsymbol{M}_t = \underset{\boldsymbol{M} \in S_+^d}{\operatorname{argmin}} \ \Delta_F(\boldsymbol{M}, \boldsymbol{M}_{t-1}) + \eta \mathcal{L}^{(t)}, \quad (4)$$

Metric Ensemble Learning for Automated Re-id

- HVIL learns a series of models $\{\boldsymbol{M}_j\}_{j=1}^{\tau}$ incrementally optimised *locally* for a set of probes with human feedback. We further propose to learn an ensemble learning model to re-id further probes without human feedback.

$$f_{ij}^{ens} = f_{\boldsymbol{x}_i^p}^{ens}(\boldsymbol{x}_j^g) = -\boldsymbol{d}_{ij}^\top \boldsymbol{W} \boldsymbol{d}_{ij}, \quad \text{s.t. } \boldsymbol{W} \in S_+^{\tau}, \quad (5)$$

4. Experiments

Human Feedback Protocol

- **Large gallery set**: Gallery contains 1000 identities.
- **Limited feedback**: Maximally 3 rounds of feedback for each probe.
- **Limited patience**: Users only verify top 5% ranked gallery images.

Human-In-The-Loop Re-Id Performance

- Supervised models were trained by an average of 3,483 cross-view images of 360 identities on CUHK03, and 7,737 images of 501 identities on Market-1501.
- Human-in-the-loop models requires maximally 3 feedbacks per probe.

Table 1: Evaluating human-in-the-loop person re-id with CMC performances.

Dataset	CUHK03 ($N_g = 1000$)				Market-1501 ($N_g = 1000$)			
Rank (%)	1	50	100	200	1	50	100	200
L2	2.9	31.1	43.2	58.2	16.1	66.6	76.6	85.0
kLFDA	5.9	47.3	60.1	75.0	21.8	85.8	**91.5**	**96.3**
XQDA	3.7	40.2	53.6	68.5	18.3	75.1	83.5	91.1
MLAPG	4.2	39.5	52.4	66.7	24.1	84.5	91.2	95.7
EMR	46.0	47.3	51.3	60.0	53.3	64.3	75.7	85.0
Rocchio	43.1	49.9	57.3	65.1	52.7	69.6	77.6	87.3
POP	46.3	55.7	64.0	74.3	56.0	72.7	80.6	86.3
HVIL (Ours)	**56.1**	**64.7**	**75.7**	**87.4**	**78.0**	**86.0**	90.3	93.4

Human Feedback Analysis

- **Boosting Rank-1 scores and pushing up true matches**

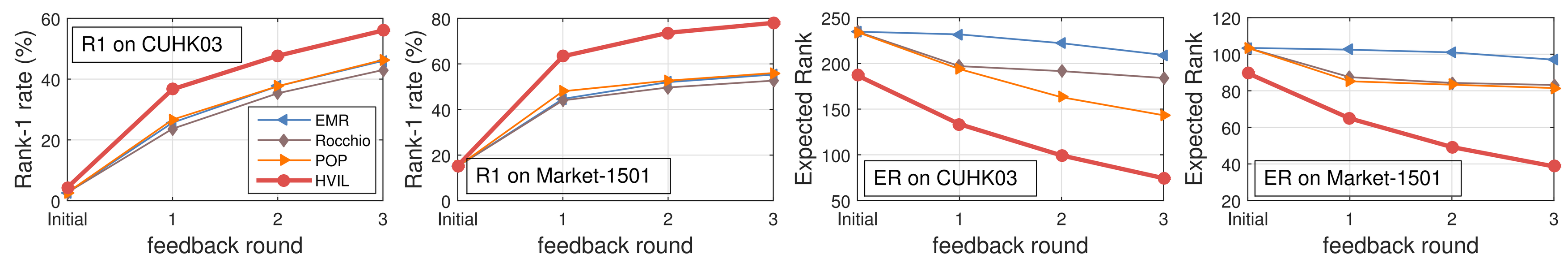


Figure 2: Comparing Rank-1 score and Expected Rank on human feedback rounds.

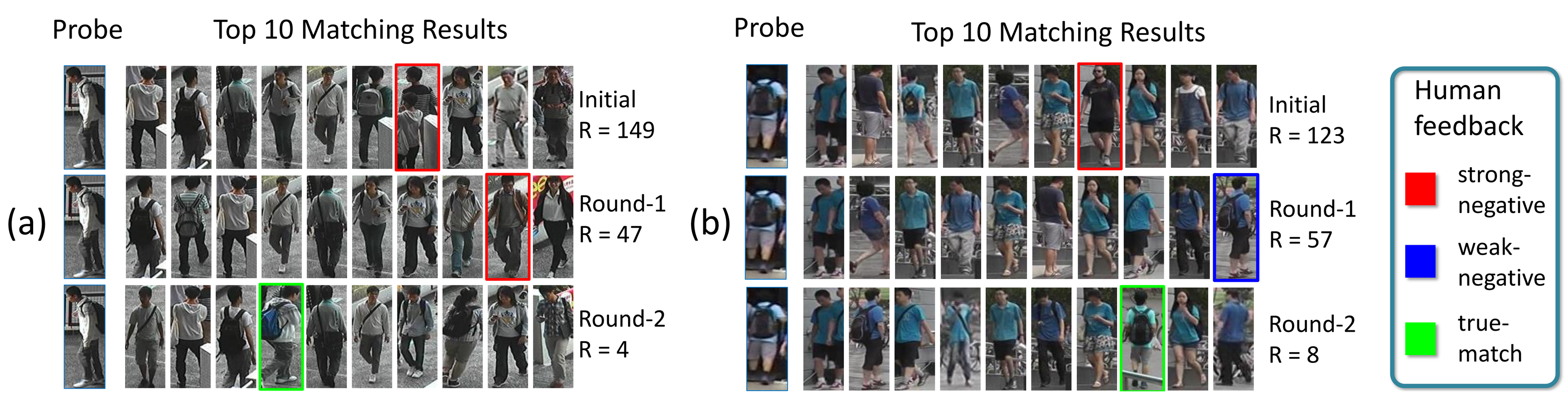
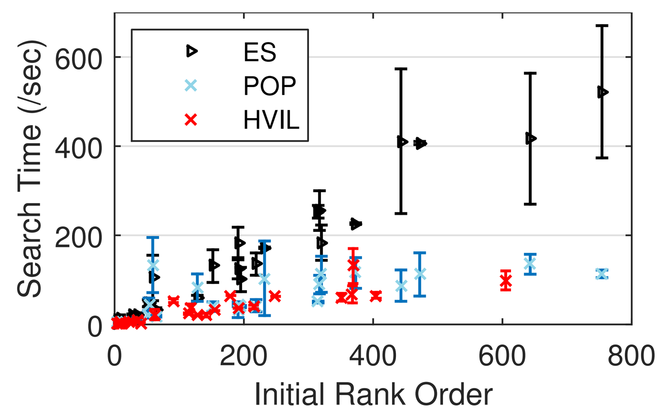


Figure 3: HVIL re-id examples on CUHK03 (a) and Market-1501 (b). The ranks of true matches before user feedback are shown.

- **Saving annotation efforts**



Dataset	CUHK03			Market-1501		
Method	HVIL	POP	ES	HVIL	POP	ES
Found-matches(%) ↑	56.1	46.3	**100**	78.0	56.0	**100**
Browsed-images ↓	**32.9**	42.1	234.1	**19.5**	38.3	108.7
Feedback ↓	**2.1**	2.3	-	**1.6**	1.9	-
Search-time(sec.) ↓	**31.7**	58.1	172.8	**28.1**	49.8	106.2

Conclusions

- A novel approach to human-in-the-loop person re-id by Human Verification Incremental Learning (HVIL) is proposed.
- HVIL model avoids the need for collecting off-line pre-labelled training data. It is able to learn cumulatively from human feedback.
- A regularised metric ensemble learning (RMEL) model is further developed to explore HVIL for automated re-id tasks when human feedback is unavailable.

References

[1] Liu, C., Loy, C.C., Gong, S., Wang, G.: POP: Person Re-identification Post-rank Optimisation, in *ICCV*, 2013.

[2] Chechik, G., Sharma, V., Shalit, U., Bengio, S.: Large scale online learning of image similarity through ranking, in *JMLR*, 2010.